M. V. JOHNS JR.

REPRESENTING FINITELY ADDITIVE INVARIANT PROBABILITIES

By

Richard A. Olshen

TECHNICAL REPORT NO. 21

March 15, 1968

PREPARED UNDER GRANT DA-ARO(D)-31-124-G726

FOR

U. S. ARMY RESEARCH OFFICE

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

REPRESENTING FINITELY ADDITIVE INVARIANT PROBABILITIES

By

Richard A. Olshen
Stanford University

1. Introduction.

Hewitt and Savage [6] have shown that finitely additive exchangeable probabilities on a product space are integral averages of power product probabilities. They prove this result as a corollary to their theorems on the countably additive case. This note adapts their technique to the study of more general invariant probabilities. From results of Farrell [4] and Choquet and Feldman ([7], Section 10) it is concluded that finitely additive invariant probabilities are averages of finitely additive ergodic probabilities.

In a countably additive context it seems necessary to impose restrictions on the Borel field being studied and on the maps used to define invariance and ergodicity. Relaxing the assumptions of one type must be balanced by strengthening those of the other (in addition to [4] and [7], see [1] and [12]). Here, however, the field of sets can be arbitrary, and the maps are assumed only to be measurable. Rather than state a host of theorems which can be proved, one particular case is proved in detail. Later on it is explained how the techniques can be applied to other problems. Several definitions of ergodicity are proposed and related to the one used. The final section contains a subjective probability interpretation of invariance and ergodicity.

Prepared with partial support of U.S. Army Research Office Grant DA-ARO(D)-31-124-G726.

2. A Representation Theorem

A homomorphism from one field of sets to another is a map which preserves finite unions, finite intersections, and complements. The notions of isomorphism and automorphism are defined in the obvious ways. A σ-homomorphism (isomorphism, automorphism) in addition preserves countable unions and intersections.

Assume Ω is a set, $\mathcal{F}$ a field of subsets of Ω, and T a 1-1 bi-$\mathcal{F}$-measurable map of Ω onto Ω. T and its powers can be viewed as automorphisms of $\mathcal{F}$. A finitely additive probability μ on $\mathcal{F}$ is said to be invariant if $\mu(A) = \mu(T^{-1}A) [= \mu(T^{n}A),\ n = \pm 1, \pm 2 \ldots]$ for all $A \in \mathcal{F}$; μ is ergodic if there do not exist $\delta > 0$ and $A_1, A_2, \ldots \in \mathcal{F}$ for which $\delta < \mu(A_n) < 1 - \delta$, $\lim\limits_{m,n\to\infty} \mu(A_n \triangle A_m) = 0$, and $\mu(A_n \triangle T^{-1} A_n) \to 0$ ($\triangle$ denotes symmetric difference). Let $\mathcal{J}(\mathcal{F})[\mathcal{E}(\mathcal{F})]$ be the set of finitely additive invariant [ergodic] probabilities on $\mathcal{F}$ and let $\mathcal{B}$ be the smallest σ-field of subsets of $\mathcal{E}(\mathcal{F})$ containing all sets of the form $\{\nu \mid \nu \in \mathcal{E}(\mathcal{F}),\ \nu(A) \leq \alpha\}$, where $0 \leq \alpha \leq 1$, and A is a fixed set in $\mathcal{F}$.

(1) THEOREM. $\mathcal{J}(\mathcal{F})$ is not empty. For each $\mu \in \mathcal{J}(\mathcal{F})$ there is a unique countably additive probability λ on $\mathcal{B}$ satisfying

$$\mu(A) = \int_{\mathcal{E}(\mathcal{F})} \nu(A)\, d\lambda(\nu) \tag{2}$$

for all $A \in \mathcal{F}$. (It is clear that, conversely, if for a finitely additive probability μ on $\mathcal{F}$ there exists a λ on $\mathcal{B}$ satisfying (2) for all $A \in \mathcal{F}$, then $\mu \in \mathcal{J}(\mathcal{F})$.)

PROOF. By Stone's representation theorem for Boolean algebras ([11], Section 8) there is an isomorphism φ of $\mathcal{F}$ onto the field of compact open subsets of some compact, totally disconnected Hausdorff

space S. Call S "the" Stone space of $\mathcal{J}$; any two such are homeomorphic. For $B \in \mathcal{O}$ define $UB = \varphi\left(T^{-1}(\varphi^{-1}B)\right)$. Of course U and its powers are automorphisms of $\mathcal{O}$. Moreover, there exists a homeomorphism ψ of S with itself for which $\psi^{-1}B = UB$ for all $B \in \mathcal{O}$ ([11], Section 11). Let $\sigma(\mathcal{O})$ be the σ-field generated by $\mathcal{O}$; it is essential to notice that $\sigma(\mathcal{O})$ is the Baire σ-field of S ([5], p. 221). Because ψ and ψ^{-1} take Baire sets to Baire sets, U can be extended uniquely to a σ-automorphism of $\sigma(\mathcal{O})$. $\mathcal{J}(\mathcal{J})$ and $\mathcal{E}(\mathcal{J})$ are in obvious 1-1 correspondence with $\mathcal{J}(\mathcal{O})$ and $\sigma(\mathcal{O})$ respectively. Each finitely additive probability η on $\mathcal{O}$ is in fact countably additive, and so admits unique extension to a countably additive probability on $\sigma(\mathcal{O})$. To see this, it is enough to demonstrate that $B_1, B_2, \ldots$ in $\mathcal{O}$ and $B_1 \supset B_2 \supset \ldots$, $\bigcap_{i=1}^{\infty} B_i = \emptyset$ imply $\eta(B_i) \to 0$ ([5], pp. 39 and 54). But this condition is trivially satisfied, for since the B's are compact and have void intersection, some B_i is already void.

Recall that if η is any countably additive probability on $\sigma(\mathcal{O})$ and for $A, B \in \sigma(\mathcal{O})$, the distance from A to B is defined to be $\eta(A \triangle B)$, then $\sigma(\mathcal{O})$ is a complete semi-metric space, and moreover that every member of the space is the limit of a Cauchy sequence of elements of $\mathcal{O}$. Easy consequences of these facts are as follows:
(a) $\mathcal{J}(\mathcal{O}) = \mathcal{J}\left(\sigma(\mathcal{O})\right)$; (b) if $\nu \in \mathcal{J}\left(\sigma(\mathcal{O})\right)$ then $\nu \in \mathcal{E}(\mathcal{O})$ iff there does not exist $B \in \sigma(\mathcal{O})$ for which $\nu(B \triangle UB) = 0$, $0 < \nu(B) < 1$. The existence of such a B implies that of a C satisfying $UC = C$, $0 < \nu(C) < 1$ -- for example, take $C = \limsup U^n B$.

An argument given by Choquet (see [7], pp. 81-2) and (a) imply $\mathcal{J}(\mathcal{J})$ is nonempty. Briefly, the set of all finite, signed, countably additive Baire measures on S can be viewed as the dual space of the Banach space

of continuous, real-valued functions on S. With those measures given the weak* topology, the maps U^n, $n = 0, \pm 1, \pm 2, \ldots$ induce continuous linear transformations of the (compact, convex) set of countably additive probabilities, $\mathcal{P}$, onto itself. And $\mathcal{I}(\sigma(\mathcal{O}))$ is the subset of $\mathcal{P}$ consisting of fixed points for the maps U^n. The Markov-Kakutani fixed point theorem ([3], p. 456) implies $\mathcal{I}(\sigma(\mathcal{O}))$ is not empty, and hence that $\mathcal{I}(\mathcal{F})$ is not empty.

In view of the existence of ψ, (a), the remark following (b), and a theorem of Farrell ([4], p. 460), for each $\mu \in \mathcal{I}(\mathcal{O})$ there exists a unique (countably additive) probability $\tilde{\lambda}$ on the σ-field of subsets of $\mathcal{E}(\mathcal{O})$ generated by sets of the form $\{\eta | \eta \in \mathcal{E}(\mathcal{O}),\ \eta(C) \leq \alpha;\ 0 \leq \alpha \leq 1,\ C \in \sigma(\mathcal{O})\}$ satisfying

$$\mu(B) = \int_{\mathcal{E}(\mathcal{O})} \eta(B)\ d\tilde{\lambda}(\eta) \tag{3}$$

for each $B \in \sigma(\mathcal{O})$. In particular (3) holds for each $B \in \mathcal{O}$. Yet for such B, $\eta(B)$ viewed as a function of $\eta \in \mathcal{E}(\mathcal{O})$ is measurable with respect to the σ-field $\tilde{\mathcal{B}}$ generated by sets of the form $\{\eta | \eta \in \mathcal{E}(\mathcal{O}),\ \eta(C) \leq \alpha;\ 0 \leq \alpha \leq 1,\ C \in \mathcal{O}\}$. Thus for $B \in \mathcal{O}$, $\tilde{\lambda}$ can be cut down to $\tilde{\mathcal{B}}$, and (3) still holds. But now (1) follows in view of the correspondence between $\tilde{\mathcal{B}}$ and $\mathcal{B}$, $\mathcal{I}(\mathcal{O})$ and $\mathcal{I}(\mathcal{F})$, and $\mathcal{E}(\mathcal{O})$ and $\mathcal{E}(\mathcal{F})$. ∎

Of course the definition of $\mathcal{E}(\mathcal{F})$ is contrived so that $\mathcal{E}(\mathcal{F})$ corresponds to countably additive probabilities on $\sigma(\mathcal{O})$ which are ergodic in the usual way. And it is implicit in (2) that $\mathcal{E}(\mathcal{F})$ is the set of extreme points of $\mathcal{I}(\mathcal{F})$. Yet it may be of interest to see what

becomes of other definitions of ergodicity when applied to $(\Omega,\mathcal{F})$. For example, $\mu \in \mathcal{S}(\mathcal{F})$ might be called ergodic if there does not exist $A \in \mathcal{F}$ for which $A = T^{-1}A$, $0 < \mu(A) < 1$. That this would have been ridiculous is clear from the following example. Let Ω be the set of bilateral sequences of 0's and 1's, $\mathcal{F}$ be the smallest field containing the cylinders, and T be the shift. It is easy to verify that Ω and $\emptyset$ are the only members of $\mathcal{F}$ in variant under T^{-1}, and so every member of $\mathcal{S}(\mathcal{F})$ would be ergodic by this definition. (With Ω given the ordinary product topology, $\mathcal{F}$ is a base of compact open sets, and so Ω is its own Stone space.)

It may seem reasonable to call $\mu \in \mathcal{S}(\mathcal{F})$ ergodic if there does not exist $B \in \mathcal{F}$ for which $\mu(B \triangle T^{-1}B) = 0$, $0 < \mu(B) < 1$. Unlike the case with $\sigma(\mathcal{O})$ and U, this definition is not identical to the one just discussed, and it is true that a μ not ergodic in this sense is not ergodic according to the definition adopted. Probabilities ergodic in this sense need not be extreme in $\mathcal{S}(\mathcal{F})$, however, as is illustrated by an example. Let $(\Omega,\mathcal{F})$ be as in the previous paragraph. Suppose μ is a measure on $\mathcal{F}$ induced by any sequence of exchangeable random variables which are neither independent nor with probability one either identically 0 or identically 1. Then it follows from deFinetti's theorem ([2], Chapter 4; [6], p. 486) that the only sets $B \in \mathcal{F}$ satisfying $\mu(B \triangle T^{-1}B) = 0$ are Ω and $\emptyset$, and that μ is not extreme in the power product probabilities on Ω, let alone in $\mathcal{S}(\mathcal{F})$.

3. Generalizations.

Apparently several times in Section 2 T^{-1} or ψ^{-1} was considered when reference to the inverse was unnecessary. The reason for this was

to keep notation consistent with that of more general problems than the one being studied. What follows is an outline of several results similar to (1), some generalizations and some not comparable. The interested reader will have no trouble filling in the details himself.

Suppose Ω is a set, $\mathcal{F}$ a field of subsets, and $\mathcal{J}$ a family of $\mathcal{F}$-measurable maps of Ω into itself. Then $T \in \mathcal{J}$ implies T^{-1} can be viewed as a homomorphism of $\mathcal{F}$ (though the same is not necessarily true of T). A finitely additive probability μ on $\mathcal{F}$ is in $\mathcal{I}(\mathcal{J})$ if $\mu(A) = \mu(T^{-1}A)$ for each $T \in \mathcal{J}$ and $A \in \mathcal{F}$; it is in $\mathcal{E}(\mathcal{J})$ if there do not exist $\delta > 0$ and $A_1, A_2, \ldots \in \mathcal{F}$ for which $\delta < \mu(A_n) < 1-\delta$, $\lim_{m,n\to\infty} \mu(A_n \triangle A_m) = 0$, and $\mu(A_n \triangle T^{-1}A_n) \to 0$ for each $T \in \mathcal{J}$.

Again the Stone space argument can be employed. If U on $\mathcal{O}$ corresponds to $T^{-1}, T \in \mathcal{J}$, there may not be a homeomorphism of the Stone space corresponding to U, but at least there is a unique continuous map ξ of S into S satisfying $UB = \xi^{-1}B$ for each $B \in \mathcal{O}$ ([11], p. 32). Hence the results of Farrell ([4], p. 460) and Choquet and Feldman ([7], pp. 82-3) apply to yield a number of theorems similar to (1). Of course, that $\mathcal{I}(\mathcal{F})$ is not empty does not follow from the Markov-Kakutani fixed point theorem unless the maps in $\mathcal{J}$ commute under composition. Yet other fixed point theorems can be employed in specific situations (see [8], Section 5; [3] p. 457).

With the definition of ergodicity given in this section, every member of $\mathcal{E}(\mathcal{J})$ is an extreme point of $\mathcal{I}(\mathcal{J})$. And the present definition is analogous to that of Phelps ([7], p. 81). Farrell's definition would have $\mu \in \mathcal{I}(\mathcal{J})$ ergodic if there does not exist $B \in \sigma(\mathcal{O})$ satisfying both $B = UB$ for each U on $\sigma(\mathcal{O})$ (corresponding to a T^{-1}) and

$0 < \nu(B) < 1$, where ν on $\sigma(\mathcal{O})$ corresponds to $\mu \in \mathcal{I}(\mathcal{F})$. For conditions that a μ ergodic in Farrell's sense be in $\mathcal{E}(\mathcal{J})$ see ([4], p. 452; [12], pp. 196-7). In general his definition is not equivalent to either definition mentioned in the last section.

It seems pointless to present a catalogue of representation theorems including results on the nonemptiness of $\mathcal{I}(\mathcal{J})$ and the uniqueness of the representation. But perhaps one striking fact deserves mention. Namely, the existence of a representation of the form (2) requires not only no assumptions on Ω and $\mathcal{J}$, but also no assumptions on the maps $\mathcal{J}$ other than their measurability. The only qualification is this. If $\mathcal{E}(\mathcal{O})$ is not weak* closed in $\mathcal{I}\big(\sigma(\mathcal{O})\big)$, the σ-field $\mathcal{B}$ in (1) must be enlarged in a manner similar to that described by Bishop and deLeeuw (see [7], pp. 31 and 83; [6], p. 481).

4. An Interpretation.

One reason for being interested in finitely additive probabilities is a sympathy with the notion of subjective probability ([2]; [10]). Thus it may seem interesting that the definitions of invariance and ergodicity used here have subjective interpretations. For purposes of illustration, assume again that Ω, $\mathcal{F}$, and T are as in the examples of Section 2. Each μ on $\mathcal{F}$ determines a law for the coordinate process $X_n(\omega) = \omega(n)$, the n-th coordinate of ω. If μ represents your beliefs about X_n, then μ is invariant for you if the probability $X_n(\omega) = 0$ (and hence the probability $X_n(\omega) = 1$) does not depend on n. μ is ergodic for you if there is an $\varepsilon > 0$ for which no pattern of 0's and 1's of finite length j $(j = 1,2,\ldots)$ has the following property. Your probability that $(X_1,X_2,\ldots,X_j)$ exhibits the pattern is strictly

between 0 and 1. But you are sure to within probability ε that $(X_1, X_2, \ldots, X_j)$ will exhibit the pattern iff $(X_2, X_3, \ldots, X_{j+1})$ will also. By contrast, the frequentistic notion of ergodicity ([9], pp. 104-5) involving averages figuring in the ergodic theorem does not fit comfortably into the framework of subjective probability (cf. [2], Chapter VI; [10], Chapter 3).

REFERENCES

[1] Cater, S. (1964). Some representation theorems for invariant probability measures. Illinois J. Math. 8, 408-418.

[2] deFinetti, B. (1937). La prevision: ses lois logiques, ses source subjectives. Ann. Inst. H. Poincaré 7 1-68. Translated by H. Kyburg and reprinted with annotations by deFinetti in Studies in Subjective Probability, H. Kyburg, Jr. and H. Smokler (editors), Wiley, 1964.

[3] Dunford, N. and Schwartz, J. T. (1958). Linear Operators, Part I. Interscience, New York.

[4] Farrell, R. H. (1962). Representation of invariant measures. Illinois J. Math. 6, 447-467.

[5] Halmos, P. R. (1950). Measure Theory. Van Nostrand, Princeton.

[6] Hewitt, E. and Savage, L. J. (1955). Symmetric measures on cartesian products. Trans. Amer. Math. Soc. 80, 470-501.

[7] Phelps, R. R. (1966). Lectures on Choquet's Theorem. Van Nostrand, Princeton.

[8] Rickert, N. W. (1967). Amenable groups and groups with the fixed point property. Trans. Amer. Math. Soc. 127, 221-232.

[9] Rosenblatt, M. (1962). Random Processes. Oxford Univ. Press, New York.

[10] Savage, L. J. (1954). The Foundations of Statistics. Wiley, New York.

[11] Sikorski, R. (1960). Boolean Algebras. Springer, Berlin.

[12] Varadarajan, V. S. (1963). Groups of automorphisms of Borel spaces. Trans. Amer. Math. Soc. 109, 191-220.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D			
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)			
1. ORIGINATING ACTIVITY (Corporate author) Stanford University, Department of Statistics, Stanford, California		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE REPRESENTING FINITELY ADDITIVE INVARIANT PROBABILITIES			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report, March, 1968			
5. AUTHOR(S) (Last name, first name, initial) Olshen, Richard A.			
6. REPORT DATE March 15, 1968	7a. TOTAL NO. OF PAGES 8		7b. NO. OF REFS 12
8a. CONTRACT OR GRANT NO. DA-ARO(D)-31-124-G726	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 21		
b. PROJECT NO.			
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)		
d.			
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.			
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Army Research Office, Durham, North Carolina		

13. ABSTRACT

Hewitt and Savage have shown that finitely additive exchangeable probabilities on a product space are integral averages of power product probabilities. They prove this result as a corollary to their theorems on the countably additive case. This note adapts their technique to the study of more general invariant probabilities. From results of Farrell and Choquet and Feldman it is concluded that finitely additive invariant probabilities are averages of finitely additive ergodic probabilities.

In a countably additive context it seems necessary to impose restrictions on the Borel field being studied and on the maps used to define invariance and ergodicity. Relaxing the assumptions of one type must be balanced by strengthening those of the other. Here, however, the field of sets can be arbitrary, and the maps are assumed only to be measurable. Rather than state a host of theorems which can be proved, one particular case is proved in detail. Later on it is explained how the techniques can be applied to other problems. Several definitions of ergodicity are proposed and related to the one used. The final section contains a subjective probability interpretation of invariance and ergodicity.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Invariant Probability Ergodic Probability Subjective Probability						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicsble, show thst optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of aervice. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the dste of the report aa day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to thia report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of thia report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
_______________________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
_______________________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
_______________________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional spsce is required, a continuation sheet shall be attached.

It is highly desirsble that the sbstract of classified reports be unclsssified. Each paragraph of the abstract shall end with an indication of the military security clsssification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitstion on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.